Evaluation of the House Price Models Using an ECM Approach: The Case of the Netherlands

Marc K. Francke, Sunčica Vujić, and Gerjan A. Vos

Ortec Finance Research Center
University of Amsterdam Business School
University of Groningen

June 27, 2009

Keywords: Bubble, Co-integration, Error-Correction Model, Long-run Equilibrium

ERES Conference, Stockholm

Purpose of this research

- Main purpose of this research:
  What is the outlook for house prices in the near future (forecasting)?
- Especially in times of a housing market crisis it is important to understand the long run price level relative to price changes in the short run.
- A long-run model approach is needed that relates house prices to fundamentals, but the model should be able to detect bubbles in the short run.
- The model assumes that prices tend to revert to the equilibrium price level in the mid term.
- Finally, the model should provide a reliable forecast of house price changes in the near future.

Research questions

- Some of the questions that we want to answer are as follows:
  - Is, or was, the Dutch housing market overvalued?
  - In which capacity can our model predict the house price developments?
  - How do housing markets react to economic growth and decline?
  - Do prices increase smoothly or unevenly during a period of adjustment to an exogenous shock?
  - Are households financially vulnerable through, for example, too high mortgage debts in comparison to disposable income?
- The research questions are investigated
  - by estimating different types of error correction models and
  - by examining the impact of different variables that can explain house price changes in the Netherlands.

What is new in the paper

- We examine the short-run and long-run price developments of the Dutch housing market in the period 1965-2009Q1.
- We evaluate (re-estimate) existing house price models for the Netherlands, which we use as a benchmark for comparison with our improved model.
- We provide a forecast of house prices until 2015, based on our own improved model.

Outline of the research

- Literature overview on house price models.
- Evaluation of estimations results of two models used in the Netherlands:
  - the OTB Research Institute 1978(1)-2000(2), based on half yearly data and
  - the Netherlands Bureau for Economic Policy Analysis (CPB), 1980-2007 with yearly data.
- Presentation of an improved error-correction model (ECM) with estimation results for an extended sample from 1965-2009Q1,
  - including ECM estimation results embedded into the unobserved components modelling approach.
- A forecast of house price changes until 2015 for three economic scenarios: recession, slow recovery and quick recovery.
- Conclusions

Type of house price models: literature overview

House price models are divided into two broad groups (De Vries and Boelhouwer, 2005) :

- demand models
  - the housing supply is fixed, and house price changes are only a function of demand variables such as housing expenses, disposable income, borrowing capacity of a consumer, etc.
- supply-and-demand models
  - both demand and supply factors are important, for example supply factors such as the housing stock and new construction.

Demand models

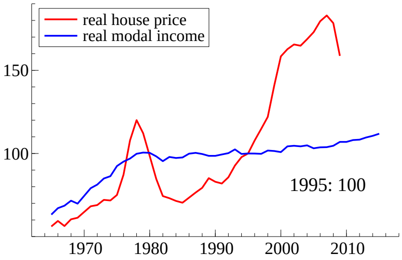
- A simple affordability model
  - The focus is on the relationship between house prices and a number of demand factors, for example, price/income ratio or mortgage-payments/income ratio. Calibration of the affordability model gives a prognosis of the house price growth in the short-run.
    - For an application, see Vos (1998, 2002), De Vries and Boelhouwer (2009).
- A dynamic error-correction model
  - The variables taken into account are: interest rate, household income, lagged house prices, and the error-correction term - the deviation from the long-run relationship. The error-correction term secures that the house prices are at their equilibrium level in the middle to long-run, which can be explained by economic fundamentals.
    - See Abraham and Hendershott (1996), Hort (1998), Malpezzi (1999), Boelhouwer (2001)

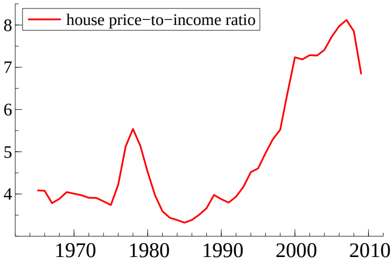
Supply-and-Demand models

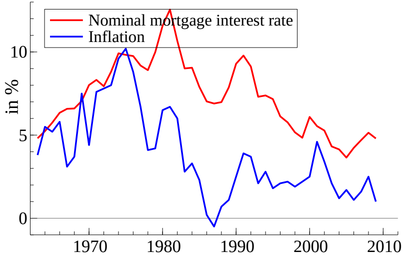
In this approach, the reaction of supply factors, like the new building developments, on demand factors can be examined (and vice versa).

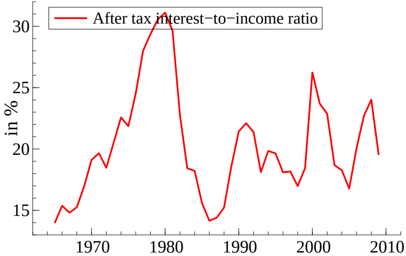
- An application is the stock-flow model which gives predictions of new construction as well as house prices through time (*see DiPasquale and Wheaton, 1994*)
- The error-correction mechanism is also applied within these supply-and-demand models, explaining house price growth in the middle to long-run (*see Verbruggen et al., 2005*).
  The variables taken into account are those which have direct effect on both house demand and supply:
  - long-run interest rate,
  - disposable household income,
  - lagged house prices,
  - housing stock,
  - number of households,
  - wealth and
  - the error-correction term (deviation from the long-run relationship).

What happened on the housing market (”65-”09Q1)?

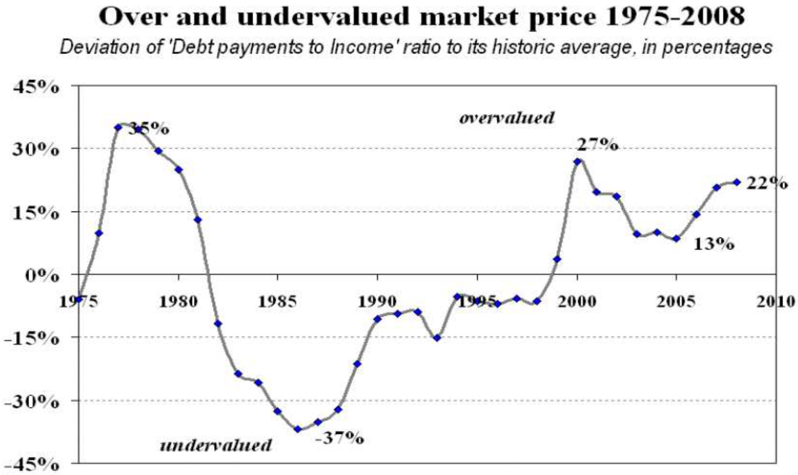








Affordability measure: over- and undervaluation



Are Median House Prices Affordable?

Based on: median house price, 100% financing of price, after tax, 5-year mortgage rate, median earnings of employees

Dutch brokerage organisation NVM 2009Q1:

- Affordable for two-earners and former owners
- Unaffordable for one-earners (only apartments are affordable)
- A more balanced judgment is possible when looking at different house types in different regions

Basic specification of Error Correction Models

h_t: log real house price
Combining both levels and differences

$$\begin{aligned} \text{Long-run:} \quad h_t^* &= \beta_1 x_{1t} + \cdots + \beta_k x_{kt}, \\ \text{Short-run:} \quad \Delta h_t &= \alpha \Delta h_{t-1} + \delta(h_{t-1} - h_{t-1}^*) + \gamma \Delta h_t^* + \varepsilon_t, \end{aligned}$$

- Long-run (levels)
  - Exogenous variables x: income, interest, wealth, construction costs, housing stock, ...
- Short-run (differences)
  - Δh_{t-1}: bubble builder: the speculative influences on the market or the market's inefficiency, α: degree of serial correlation
  - $(h_t - h_t^*)$ bubble burster: error correction term, the deviation from the long term equilibrium, δ: degree of mean reversion
    - $(h_t - h_t^*) > 0$: overvaluation
    - $(h_t - h_t^*) < 0$: undervaluation
  - γ: contemporaneous adjustments of prices to current shocks

Characteristics co-integrated Error Correction Models

- h_t is non-stationary, I(1): integrated of order 1 (no time invariant first and second moment)
- If $h_t - h_t^* = h_t - x_t\beta$ for some β is
  - stationary, then a co-integrating relation: $x_t\beta$ represents the long-run equilibrium relationship
  - non-stationary, then spurious regression results: the ECM is not valid, the usual statistics (standard errors, R^2, etc.) do not have their common interpretation
- Extensions of ECM
  - More lags of Δh_t and Δh_t^* can be included
  - Asymmetric ECM: different coefficients for positive and negative values of error correction term
- Estimation of ECM
  - Dynamic Linear Regression Model (PcGive, Hendry)
  - Test the null hypothesis of no co-integration by residual based augmented Dickey-Fuller test

CPB Model specification

- The Netherlands Bureau for Economic Policy Analysis (CPB) Verbruggen et al. (2005); Kranendonk and Verbruggen (2008)
- Variables used:
  - House price index (Kadaster)
  - Disposable labour income (aggregate)
  - Long-term interest rate (10-year government securities)
  - Wealth indicator
  - Housing stock (end of the year)
  - Consumer price index
- Yearly data: 1980 – 2007

CPB Long and Short Run Re-Estimation Results

Long-run h_t

Variable	Coefficient	Std. Error	t-value
Constant	−6.5986	1.2730	−5.18
y_t	1.5336	0.2653	5.78
l_t^r	−5.944	1.6900	−3.52
w_t	1.6320	0.4201	3.89
s_t	−2.8298	0.6032	−4.69
Sigma = 0.0703		RSS = 0.1136	
$R^2 = 0.9705$		DW = 1.11	

Short-run Δh_t

Variable	Coefficient	Std. Error	t-value
Δy_t^a	1.4386	0.2225	6.46
Δl_t^a	−6.3515	1.3530	−4.70
$\Delta \nabla P_t$	1.1015	0.8375	1.32
Δs_t	−2.0639	0.5536	−3.70
d_{2000}	0.1398	0.0358	3.90
ecm_{t-1}	−0.2177	0.1852	−1.18
ecm_{t-1}^+	0.3238	0.2905	1.11
Sigma = 0.0339		DW = 1.42	

CPB Long and Short Run Re-Estimation Results

- Long-run
  - No evidence for co-integration (augmented DF-test)
  - Overvaluation:
    - 2004: +14%
    - 2007: 0%
- Short-run
  - Ad hoc dummy variable for $t = 2000$
  - No lagged real house price changes
  - No evidence for asymmetric error correction mechanism

OTB Model Specification

- Half-Yearly data: 1978-2000
- Variables used:
  - (Real) Mortgage interest (I)
  - (Real) Household income (Y)
  - Seasonal effect (d)
- Model in percentage changes (∇) (not in log differences)
- "Error correction term": $(IIR_t - \gamma_0)$ where IIR_t is the after tax (F) Interest-to-Income ratio, defined by

$$IIR_t = \frac{H_t I_t (1 - F)}{Y_t},$$

- γ_0 is the long-run Interest-to-Income ratio
- Model (Boelhouwer et al, 2001, De Vries and Boelhouwer, 2009)

$$\nabla H_t = \gamma_1 \nabla H_{t-1} + \gamma_2 (IIR_{t-2} - \gamma_0) + \gamma_3 d_t + \gamma_4 \nabla Y_t + \gamma_5 \nabla I_t^r + \varepsilon_t$$

OTB Re-Estimation Results 1978-2000 (Half-Yearly)

Variable	Coefficient	Std. Error	t-value
Constant	5.2626	1.6240	3.24
∇H_{t-1}	0.5574	0.0863	6.46
IIR_{t-2}	-0.1926	0.0583	-3.30
d_t	1.3522	0.3001	4.51
∇Y_t	0.5657	0.2520	2.24
$\gamma_0 = 5.2626/0.1926 = 27.33$			
Sigma = 1.9902		RSS = 158.4371	
$R^2 = 0.8230$		DW = 1.71	

- IIR_t is not stationary (however bounded between 0 and 100)
- OTB model is a restricted version of a general ECM

$$\Delta h_t = \gamma_1 \Delta h_{t-1} + \gamma_2 (h - (1y - 1i) - \gamma_0^*)_{t-2} + \gamma_3 d_t + \gamma_4 \Delta y_t + \gamma_5 \Delta I_t^r + \varepsilon_t$$

using $IIR_t \approx h_t + i_t + \ln(1 - F) - y_t - 1$ and $\Delta x_t \approx \nabla x_t$

Our improved Unobserved Component ECM

- Extended sample: 1967 - 2009Q1
- Used variables
  - Real log house price: I(1)
  - Log real modal labour income per employee: I(0)
  - Mortgage interest rate minus inflation (not in logs): I(1)
  - Linear trend (long-run relation)
- However, ECM term is non-stationary: no co-integration
- We replace the linear trend with a non-stationary trend

$$\Delta h_t = \alpha \Delta h_{t-1} + \delta(h_{t-1} - \mu_t - x_{t-1}\beta) + \gamma \Delta x_t + \varepsilon_t,$$

  where μ_t denotes the trend component (Harvey, 1989)
- In case of a random walk, the trend component is given by:

$$\mu_{t+1} = \mu_t + \eta_t$$

ECM with a Random Walk Estimation Results

- Estimation results

$$\Delta h_t = 0.4726\Delta h_{t-1} - 0.3776 ECM_{t-1} + 0.01534\Delta IM_{t-1} - 0.5372\Delta y_{t-1}$$

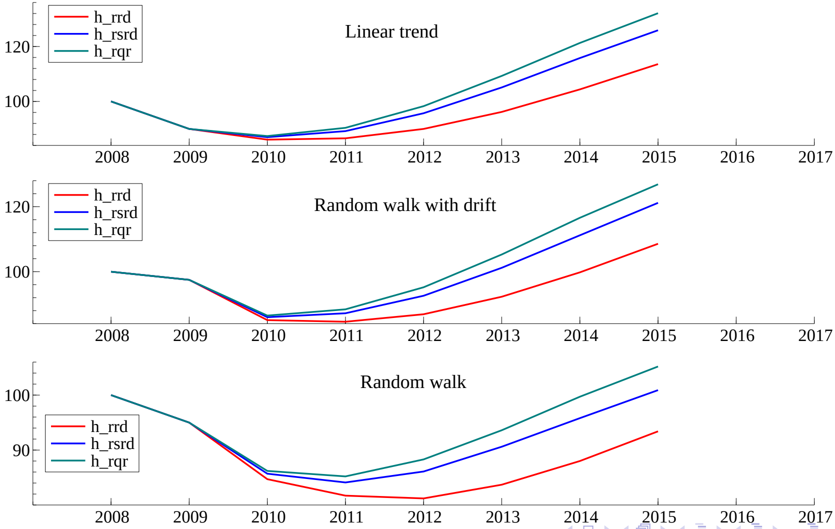
where

$$ECM_t = h_t - 1.6766y_t + 0.0681IM_t - \mu_{t+1}.$$

- Standard error random walk: 0.030
  (different from 0: alternative test for co-integration)
- Long-term relationship: increase in period 1967 - 2009Q1

Real Income	0.744
Real Interest	-0.022
Random Walk	0.227
Sum	0.949
Real House Price Increase	1.035

Forecasts - Price Index (nominal values)



Conclusions to models

- Compared to other Dutch research institutes this research presents estimation results for a much longer sample (1967-2009).
    - It is important to stress that our approach nicely estimates a full cycle of house price movements from 1970 to 1980, the period not analysed by the other research institutes.
- Other differences with the CPB specification are that we also include lagged changes in log real house prices.
- With respect to the OTB approach, our model differs in that we present results based on an unrestricted model.

Conclusions to valuation

- All models estimate the overvaluation of the Dutch house prices differently.
  - Based on a simple affordability approach analysis it can be concluded that house prices in 2007-2008 were around 18% overvalued compared to the long-run interest payments-to-income ratio.
  - Due to a price fall of around 8% and the lower mortgage interest rate in the first half year of 2009 average affordability is in line with its historic average (median price level is not overvalued).
  - From the CPB long-run relation it can be concluded that in 2007 the overvaluation was approximately 0, whereas it was +14% in 2004.
  - Our ECM model with linear trend estimates that the Dutch house prices were severely undervalued in 1975 (-35.6%), followed by a period of extreme overvaluation in 1978 (42.9%). During the 2000s, house prices in the Netherlands were also overvalued (2006: 11.9%; 2007: 11.5%; 2008: 2.4%).
    Note that part of the linear trend can capture overvaluation as well.

Conclusions to our improved model

- The linear trend and random walk (with drift) component indicates that a substantial part of house prices has not been explained by fundamental economic factors.
  - If we interpret the random walk (with drift) term together with the error-correction term as a way measuring overvaluation in the housing market, they indicate that Dutch house prices were substantially overvalued in the last decade.
  - In contrast, if overvaluation is measured by the error-correction term only, house prices were moderately above the long-run equilibrium value.
  - One can argue that a part of the random walk component captures some omitted variables in the long-term relationship. In that case only a fraction of the random walk (with drift) component can be interpreted as overvaluation.
- Our preferred model is ECM with random walk (lowest standard error). This model is the most “pessimistic” one, considering the forecasting scenarios and the overvaluation estimate.

Conclusions to forecasting

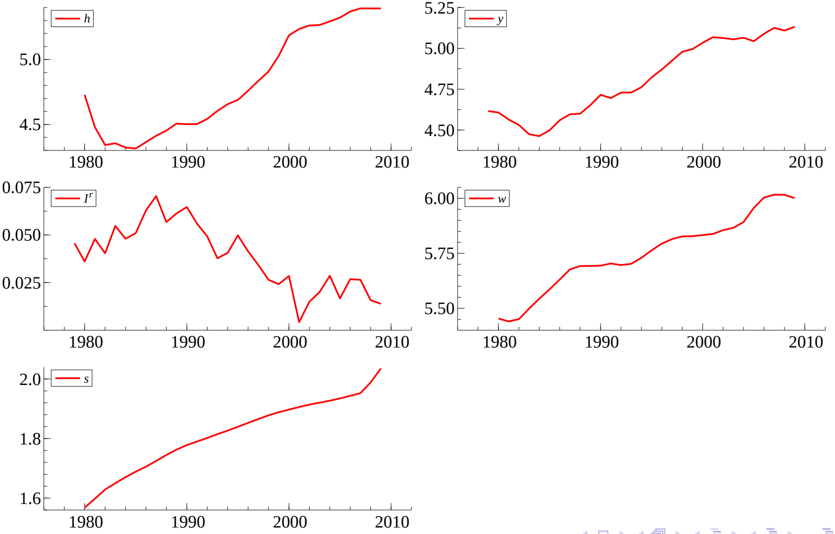
- The current financial crisis in the Netherlands did not start with problems in the residential property market, but on the contrary, the global financial and economic crisis has affected the housing market.
- Forecasting house prices with this model shows a recovery of prices to the level of 2008 no sooner than 2015 in all scenarios, except for the recession scenario.

Further Research

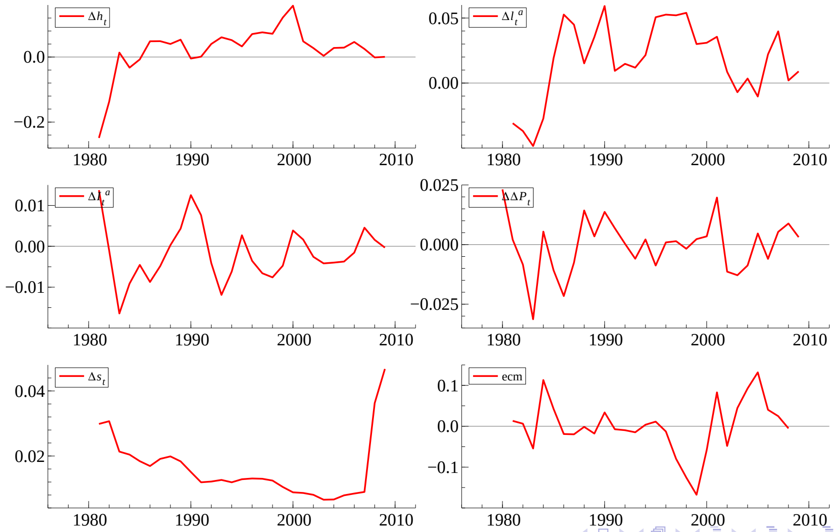
Further research should encompass several extensions of the current model.

- The real interest rate could also be based on expected inflation, in order to account for real user costs.
- Among the set of explanatory variables we would also like to include housing stock and construction costs (the supply side of the market).
- We would like to improve the household income data taking into account the transition from a one-earner into a two-earner economy since 1985.
- We want to present estimation results on a more disaggregate level, such as regions or the largest cities in the Netherlands.

CPB Long Run Series



CPB Short Run Series



CPB Long and Short Run Relationships

- The long-term equation

$$h_t = \alpha_0 + \alpha_1 y_t + \alpha_2 l_t^r + \alpha_3 w_t + \alpha_4 s_t + \varepsilon_t, \text{ for } t = 1980, \ldots, 2007.$$

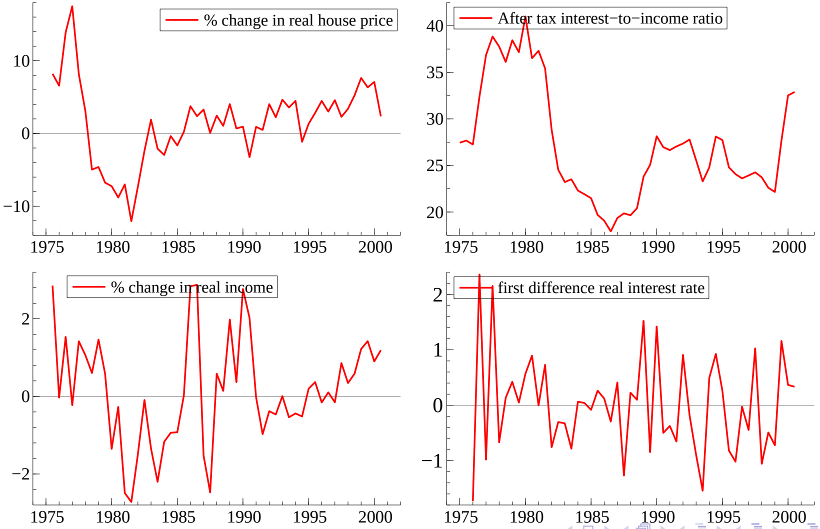
- The short-term equation

$$\begin{aligned} \Delta h_t &= \beta_1 \Delta y_t^a + \beta_2 \Delta l_t^a + \beta_3 \Delta \nabla P_t + \beta_4 \Delta s_t + \beta_5 d_{2000} \\ &+ \beta_6 ecm_{t-1} + \beta_7 ecm_{t-1}^{+} + \varepsilon_t \end{aligned}$$

where

$$\begin{aligned} \Delta h_t &= h_t - h_{t-1} \\ \Delta y_t^a &= 0.65 \Delta y_t + 0.35 \Delta y_{t-1} \\ \Delta l_t^a &= 0.5 \Delta l_t + 0.5 \Delta l_{t-1} \\ \Delta \nabla P_t &= \nabla P_t - \nabla P_{t-1} \\ \Delta s_t &= s_t - s_{t-1} \end{aligned}$$

OTB Series (Percentage Changes)



ECM with a Linear Trend

- General-to-specific modelling approach

$$h_t = \alpha_1 h_{t-1} + \alpha_2 h_{t-2} + \sum_{i=1}^{k} \beta_{i0} x_{it} + \sum_{i=1}^{k} \beta_{i1} x_{i,t-1} + \sum_{i=1}^{k} \beta_{i2} x_{i,t-2} + \varepsilon_t .$$

- The long-run equilibrium (substitute $h_t = h^*$ and $x_{it} = x_i^*$)

$$h^* = \frac{1}{1 - \alpha_1 - \alpha_2} \sum_{i=1}^{k} x_i^* (\beta_{i0} + \beta_{i1} + \beta_{i2})$$

ECM with a Linear Trend Estimation Results

Variable	Coefficient	Std. Error	t-value
h_{t-1}	1.2952	0.1189	10.90
h_{t-2}	-0.6100	0.1037	-5.88
y_{t-2}	0.3002	0.1110	2.70
IM_{t-1}	-0.0122	0.0054	-2.24
IM_{t-2}	-0.0149	0.0062	-2.40
Trend	0.0060	0.0015	4.08
Constant	-0.0428	0.0417	-1.03

Sigma = 0.0439	RSS = 0.0692
$R^2 = 0.9872$	$F(6, 36) = 463.9(0.0000)$
Log-likelihood = 77.2615	DW = 1.88
No. of observations = 43	No. of parameters = 7
$\bar{h}_t = -0.0126$	$\sigma^2_{h_t} = 0.1261$

Static Long-Run Estimation Results

Variable	Coefficient	Std. Error	t-value
y_t	0.9534	0.3156	3.02
IM_t	-0.0859	0.0131	-6.57
Trend	0.0190	0.0032	6.04
Constant	-0.1360	0.1274	-1.07
Long-run sigma = 0.1393			
$ECM = h_t - 0.9534 y_t + 0.0859 IM_t - 0.0190 Trend + 0.1360$			

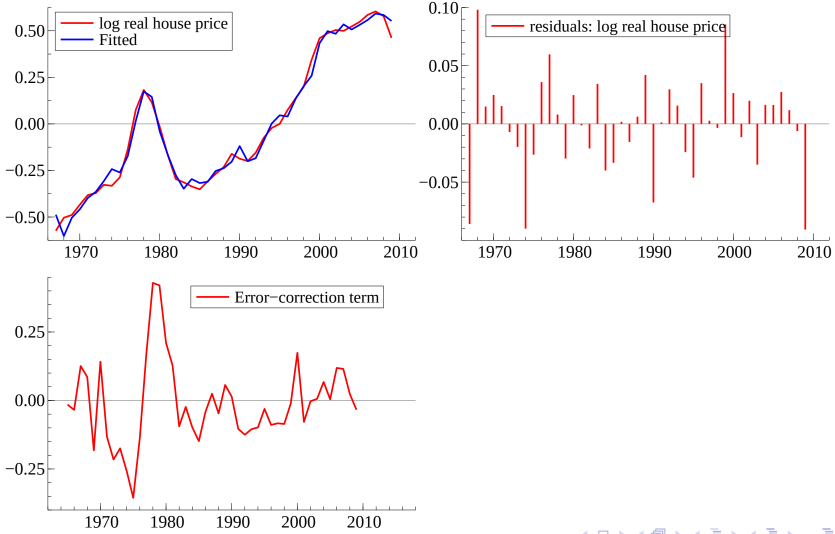
- Expressed in terms of the ECM model specification, we get:

$$\Delta h_t = 0.6142 \Delta h_{t-1} - 0.3149 ECM_{t-1} - 0.3002 \Delta y_{t-1} + 0.0149 \Delta IM_{t-1}$$

Interpretation of the Estimation Results

- The marginal long-run real income elasticity is close to one.
- Linear trend included to capture absence of other important variables, like demographics and house supply.
- In comparison to the CPB and OTB models, our sample is much longer (1967 to 2009Q1):
  - CPB sample: 1980 – 2007.
  - OTB sample: 1978 – 2000.
  - We also model house price movements from 1970 to 1980.
- However, the ECM term is *not* stationary!!
- We replace the linear trend with a (non-stationary) random walk.

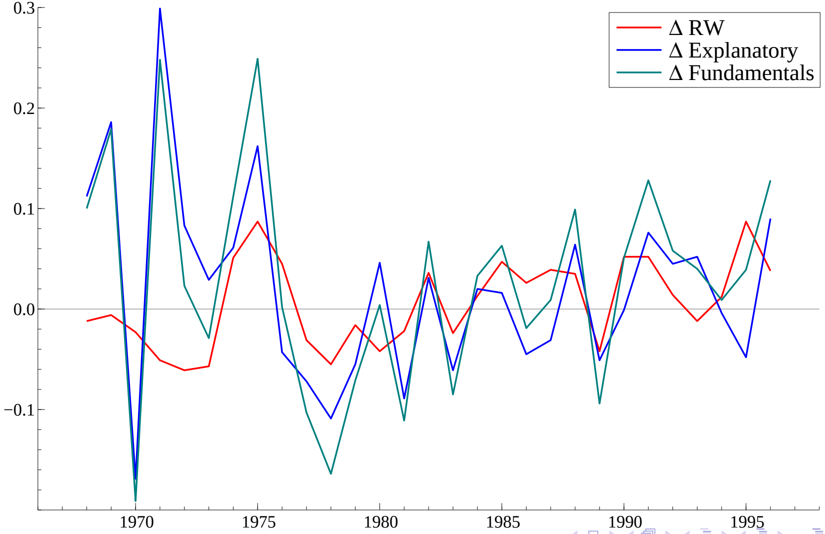
ECM with a Linear Trend, 1965-2009Q1



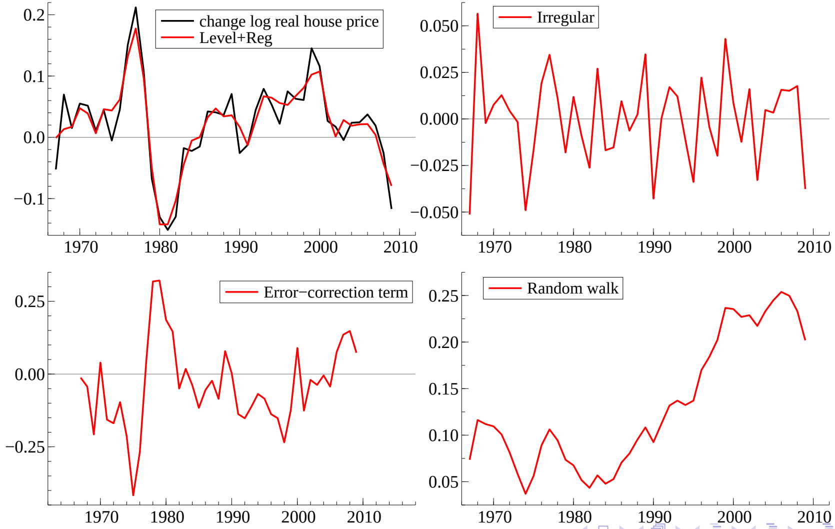
ECM with a Random Walk Estimation Results

Variable	Coefficient	Std. Error	t-value
Δh_{t-1}	0.4726	0.1391	3.3986
ΔIM_{t-1}	0.0153	0.0065	2.3448
Δy_{t-1}	-0.5372	0.3739	-1.4367
h_{t-1}	-0.377	0.0943	-4.0035
IM_{t-1}	-0.0257	0.0089	-2.8870
y_{t-1}	0.6330	0.2952	2.1445
$ECM = h_t - 1.6766y_t + 0.0681IM_t - \mu_{t+1}$			
Disturbances	Variance	Std. Error	
Level	0.0009	0.0300	
Irregular	0.0011	0.0330	
State vector	Value	p-value	
Level	0.2020	0.0024	
Std. error = 0.0473		Log-likelihood = 107.2840	
No. of observations = 43		p.e.v. = 0.0022	
$R^2 = 0.6418$		DW = 1.7405	

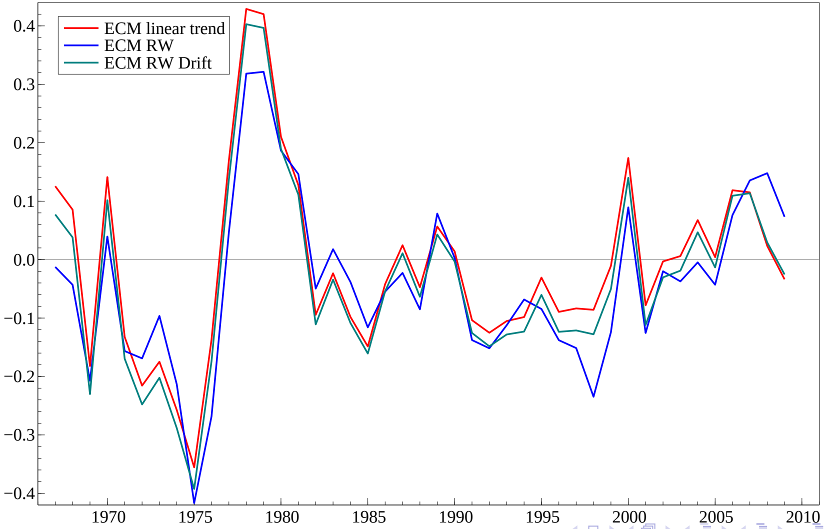
ECM with a Random Walk, 1965-2009Q1



ECM with a Random Walk, 1965-2009Q1



Comparison of Error-Correction Models



Overvaluation of House Prices in NL

	Random walk		Random walk with drift	
Year	*ECM*	*ECM* $+ \mu_{t+1}$	*ECM*	*ECM* $+ \mu_{t+1}$
2000	0.0895	0.3249	0.1400	0.3335
2001	−0.1256	0.1015	−0.1118	0.0862
2002	−0.0200	0.2088	−0.0304	0.1727
2003	−0.0373	0.1801	−0.0189	0.1885
2004	−0.0048	0.2283	0.0467	0.2595
2005	−0.0430	0.2018	−0.0133	0.2045
2006	0.0761	0.3299	0.1092	0.3316
2007	0.1355	0.3852	0.1136	0.3399
2008	0.1479	0.3811	0.0294	0.2592
2009*Q*1	0.0733	0.2752	−0.0255	0.2079

Overvaluation:

- ECM plus random walk: 18% to 38%
- ECM only: −13% to 15%

Forecasts (2010-2015)

- Three scenarios
- Forecast period: 2010 - 2015

	Recession			Slow Recovery			Quick Recovery		
Year	dP	IM	Y	dP	IM	Y	dP	IM	Y
2010	0.00	4.50	32,000	1.0	4.70	32,400	1.5	4.80	32,500
2011	0.50	4.60	32,500	1.5	4.90	33,000	1.8	4.90	33,200
2012	1.00	4.80	32,900	1.8	4.90	33,600	2.2	5.00	34,000
2013	1.20	4.80	33,700	1.9	5.00	34,500	2.2	5.00	35,000
2014	1.50	4.90	34,500	2.0	5.00	35,000	2.3	5.20	36,200
2015	1.70	4.90	35,500	2.2	5.00	36,600	2.5	5.20	37,500

Forecasts (2010-2015)

Year	Recession			Slow Recovery			Quick Recovery		
Model	Linear	RWD	RW	Linear	RWD	RW	Linear	RWD	RW
2008	*100*	*100*	*100*	*100*	*100*	*100*	*100*	*100*	*100*
2009	*90.0*	*97.5*	*95.0*	*90.0*	*97.5*	*95.0*	*90.0*	*97.5*	*95.0*
2010	86.1	85.1	84.7	87.0	86.0	85.7	87.4	86.5	86.2
2011	86.6	84.6	81.7	89.2	87.2	84.1	90.4	88.4	85.2
2012	90.0	86.9	81.2	95.7	92.6	86.1	98.3	95.2	88.3
2013	96.2	92.3	83.7	105.1	101.2	90.6	109.3	105.3	93.6
2014	104.4	99.8	88.0	115.8	111.2	95.8	121.3	116.6	99.7
2015	113.6	108.6	93.4	125.9	121.2	100.9	132.1	126.9	105.2
	RW = Random walk; RWD = Random walk with drift								